Food Control 103 (2019) 1–8

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Validation protocol for commercial sterility testing methods

Benjamin Diep*, Julie Moulin, Viktoria Bastic-Schmid, Thierry Putallaz, Johan Gimonet, Antonio Deban Valles, Adrianne Klijn

Department of Food Safety Research, Nestlé Research, CH-1000, Lausanne, Switzerland

ARTICLE INFO

Keywords:
Spoilage
Alternative methods
UHT
Aseptic filling products

ABSTRACT

Thermal processing technology has been widely applied for the preservation of food. Initially used for canned foods, thermal processes have since been extended to a large range of foods. Ready-to-eat products processed at Ultra-High Temperature (UHT) and aseptically filled are extensively consumed because of their convenience. Sterilisation and aseptic filling are critical steps, and food business operators have to verify their efficacy by demonstrating commercial sterility. Methods commonly used to demonstrate commercial sterility also originate from the canning industry and are both cumbersome and time consuming. Several alternative methods are available, but they do not have official validation status since standard validation protocols, such as ISO 16140–2 and AOAC guidelines cannot be applied due to differences in the testing procedure. We propose a validation protocol based on inclusivity and limit of detection LOD_{95} as performance criteria. The traditional direct streaking and six alternative methods were assessed to demonstrate the relevance of the protocol; three methods were based on cellular metabolism during microbial growth (CO_2 production, O_2 consumption), one was based on cell count by flow cytometer and two methods were based on cellular ATP activity. Nine food items including challenging matrices (high pH, high fat contents) were tested with sporeforming and non-sporeforming microorganims. Inclusivity results show that all the methods could detect a large range of microorganisms provided appropriate culture media were used. The LOD_{95} results indicated that methods based on cellular metabolisms were very sensitive ($LOD_{95} < 1 \log_{10}$ CFU/mL) compared to cells counts and ATP-based methods ($LOD_{95} > 3 \log_{10}$ CFU/mL). This is the first study proposing relevant performance criteria to validate alternative commercial sterility methods. The outcomes allow the end user to select a right method according to their requirements.

1. Introduction

Thermal processing technologies are widely used in food manufacturing to render food safe from pathogens and spoilage microorganisms, without affecting the product organoleptic and nutritional quality. Because absolute sterility cannot be achieved without drastically damaging food integrity, thermal processed products must meet the commercial sterility criterion before being placed on the market. Commercial sterility refers to the absence of microorganisms capable of growing in the food at normal non-refrigerated conditions at which the food is likely to be held during distribution and storage (Codex Alimentarius; CAC/RCP 40–1993 (Anonymous, 1993)). This is achieved by robust process design based on Hazard Analysis Critical Control Point (HACCP) principles (Sperber, 2005) and the verification of the implemented controls measures (Ropkins & Beck, 2000) via commercial sterility testing. In the past, thermal processed products referred mostly to canned products. Today, thermal processes in combination with other preservation processes are applied to a wide range of products and packaging formats (Sanchez-Madrid, 2003; Datta, 2018). Ultra-High Temperature (UHT) processed and aseptically filled products, described in this paper, are extensively consumed throughout the world. They include products such as liquid dairy products, baby foods, desserts, sauces, soups, fruit juices and soft drinks. After a high heat treatment (135 °C–150 °C) for a very short time (3–5 s), they are aseptically filled into a variety of packaging format, including cardboard cartons, plastic bottles and foil pouches. UHT processing and aseptic filling are designed to reach commercial sterility, but microbial defects can occur at each stage of the production. Insufficient heat treatment or high contamination of raw materials may lead to food spoilage by sporeforming microorganisms. Inappropriate cleaning and sterilisation of processing and filling equipment or faulty packaging can lead to post-process contamination (Pujol, Albert, Johnson, & Membré, 2013) by thermophilic and mesophilic sporeforming microorganisms and vegetative microorganisms (Eneroth, Ahrneh, & Molin, 2000). The

* Corresponding author.
E-mail address: benjamin.diep@rdls.nestle.com (B. Diep).

https://doi.org/10.1016/j.foodcont.2019.03.029
Received 30 January 2019; Received in revised form 21 March 2019; Accepted 22 March 2019
Available online 23 March 2019

Table 1
Performance characteristics to validate sterility testing methods.

	Selectivity	Sensitivity
Performance charactericties	Inclusivity	LOD_{95}
In the scope of this study	Sporeforming and non sporeforming microorganism	Sensitivity of each method after products incubation. No stressed cells, no spores are required
Out of the scope of this study	Exclusivity	Products incubation time (7–13 days at 30 °C)

objective of commercial sterility testing is then to verify manufacturing procedures and that products are indeed commercially sterile. For UHT processed and aseptically filled products, food operators often refer to the protocols used for canned food testing (AFNOR NF V08-401 (Anonymous, 1997), FDA BAM (Landry, Schwab, & Lancette, 2001)). In these two national standards, a first step is the incubation of the products in its final packaging for seven to 13 day at 30 °C (Krumm, Lattuada, Johnston, Eye, & Green, 1998, chap. 10; Landry et al., 2001; Mostert & Jooste, 2005) to allow surviving spores or contaminating microorganisms to recover and grow to a detectable level. Extra incubation at 55 °C for five to seven days is also used when the products are intended to be stored at higher temperatures (> 40 °C). A second step focuses on the detection of the viable microorganisms that grew in the product using macroscopic and microscopic examinations, pH measurement and direct streaking. Since many years, several commercial testing solutions have been proposed to reduce the time-to-result (which can go up to 10 days) as well as to shorten the product incubation time. Those methods were validated either as a quantitative method (Bolzoni, Marcolini, Delle Donne, Appicciafuoco, & Ferrini, 2015) or as a semi-quantitative method (Fernandes, Carey, Hynes, & Papkovsky, 2013; Mozola et al., 2013), but their performance as a qualitative method to assess commercial sterility has not been demonstrated. The reason for this is that existing validation protocols such as ISO 16140–2 (Anonymous, 2016) or the Association of Analytical Communities AOAC guidelines (Feldsine, Abeyta, & Andrews, 2002) cannot be applied as such. Typical performance characteristics are inclusivity/exclusivity, the Limit Of Detection (LOD_{50}) and a study to compare the sensitivity of the reference and alternative method with various contaminated food items. Inclusivity/exclusivity determines the selectivity of a method by testing a set of target microorganisms and non-target microorganisms. However, this criterion is not mandatory for a method targeting a large group of microorganisms such as total viable count or yeast and mold as the method is supposed to cover the complete range. For commercial sterility methods, only the microorganisms that are able to grow in the food item should be taken into consideration. Therefore, inclusivity is an important criterion and a set of microorganism relevant to the food item must be selected. Exclusivity is not an important criterion as microorganisms that are not relevant will not grow in the product and therefore not detected. Typical qualitative methods validated using the ISO 16140–2 protocol include an enrichment step to ensure the growth of target microorganism (Fig. 5) and the complete procedure (enrichment and detection) is validated. As these methods target very low levels of detection, food items are inoculated at fractional level and the LOD_{50} which, represents the concentration at which the microorganism is detected in 50% of cases, is used. For commercial sterility methods, recovery and multiplication of microorganisms take place in the food item itself, leading to an unknown microorganism concentration before the detection step. Therefore, to quantify the sensitivity of commercial sterility methods, defined here as the ability of the method to detect the microorganism after the product incubation time, we propose a protocol that focusses only on the detection step. Food items are inoculated with different concentrations of microorganism, but not at fractional level, and thus the LOD_{95} which determine the concentration of microorganism detected in 95% of cases is used. The sensitivity study as described in ISO 16140–2, is not applied as we do not include the product incubation step and do not use fractional level for the reason mentioned above. To demonstrate the relevance of our protocol, six alternative methods and the direct streaking method were assessed. The alternative methods included three methods based on the detection of metabolites during microbial growth in a specific broth (CO_2 detection, O_2 consumption), three methods were based on rapid measurement of ATP activity (ATP-based methods) and cell counting (flow cytometer). Nine UHT processed and aseptically filled food items including high and low acid food items were inoculated with and a range of sporeforming and non-sporeforming microorganisms that mimic different contamination scenarios. Table 1 summarizes the performance characteristics of the validation protocol. Assessing the method performance allows the food business operator to decide on a method or a combination of methods aligned to their requirements, knowing that a highly sensitive method may permit (1) shorter product incubation time, (2) detection of slow grow microorganisms.

2. Materials and methods

2.1. 1. food items

Selected food items, including complex matrices such as fruit particles and food items with high fat content, which may interfere with the equipment performance or cell growth are shown in Table 2.

2.2. 2. microorganism strains and strain activation

A panel of relevant spoilage organisms was selected (Table 3) and included both sporeforming and non-sporeforming microorganisms, the latter being relevant in the case of post-process contamination. In line

Table 2
Food items.

Food item	pH	Food item description	Packaging
Non-hydrolysed infant formula	pH 6.77	Dairy drink	Glass bottle
Whey protein hydrolysed infant formula	pH 6.71	Dairy drink	Glass bottle
High protein-vanilla flavor dairy	pH 6.89	Dairy drink	Plastic bottle
Evaporated milk	pH 6.40	Dairy drink with 60% water removed	Cardboard carton
Chocolate dairy	pH 6.75	Dairy drink	Plastic bottle
Banana-Apple-Strawberry puree	pH 3.9	High acid puree	Glass bottle
High energetic drink	pH 4.13	High acid dairy drink	Plastic bottle
Carrot puree	pH 5.1	Low acid puree	Glass bottle
Béchamel sauce	pH 6.53	High fat content sauce	Cardboard carton

Table 3
Inclusivity results.

Strain	Strain source	Greenlight	Bact/Alert	Soleris	Bactiflow	MLS	Cellscan
Bacillus mycoides B 162[a]	Milk	+	+	+	+	+	+
Bacillus pumilus B 348	Tomato	+	+	+	+	+	+
Bacillus firmus ATCC 14575[b]	Unknown	+	+	+	+	+	+
Bacillus coagulans ATCC 7050	Evaporated milk	+	–	+	+	+	+
Bacillus sporothermodurans B 515	Evaporated UHT milk	+	+	+	+	+	+
Bacillus subtilis ATCC 14807	Unknown	+	+	+	+	+	+
Bacillus licheniformis ATCC 9789	Unknown	+	+	+	+	+	+
Bacillus flexus B 582	UHT cocoa drink	+	+	+	+	+	+
Geobacillus stearothermophilus B 574	Whey powder	–	–	+	–	–	–
Lysinibacillus sphaericus B 400	Sponge cake	+	+	+	+	+	+
Staphylococcus aureus STA 057[a]	Ice cream	+	+	+	+	+	+
Staphylococcus warneri STA 097	Frozen lactic acid starter	+	+	+	+	+	+
Micrococcus luteus DSM 20030[c]	Unknown	+	+	–	+	+	+
Enterococcus faecalis ATCC 2320	Unknown	+	+	+	+	+	+
Pseudomonas aeruginosa NE 49a	Unknown	+	+	+	+	+	+
Pseudomonas fluorescens ATCC 13525	Unknown	+	+	+	+	+	+
Enterobacter cloacae ATCC 23355	Unknown	+	+	+	+	+	+
Saccharomyces cerevisiae Y 023[a]	Cheese	+	+	+	+	+	+
Zygosaccharomyces bailii Y 063	Mayonnaise	+	+	+	+	+	+
Candida albicans Y 216	Ice tea	+	+	+	+	+	+
Candida magnoliae Y 411	Jam	+	+	+	+	+	+
Byssochlamys fulva M 233[a]	Grapes	+	+	+	+	+	+
Aspergillus carbonarius M 325	Apple	+	+	+	+	–	+
Eurotium amstelodami M 608	Petfood environment	+	+	–	+	–	–
Penicillium waksmanii M 586	Soft drink	+	+	+	+	+	+
Lactobacillus plantarum NCC 1594[a]	Unknown	+	+	+	+	+	+
Lactobacillus brevis NCC 2792	Unknown	+	+	+	+	+	+
Leuconostoc mesenteroides NCC 2825	Unknown	–	+	+	+	–	+

[a] Internal reference: B: Bacillus, STA: *Staphylococacceae/Micrococacceae,* NE: Non-Enterobacteriaceae, Y: Yeast, M: Mold, NCC: Nestlé Culture Collection.
[b] American Type Culture Collection.
[c] German collection of microorganisms and cells cultures.

with the method principles, only aerobic microorganisms were tested. Each strain was transferred from frozen stock culture and grown under optimum conditions. *Bacillus*, *Staphylococcus*, *Micrococcus*, Enterobacteriaceae and *Pseudomonas* were incubated in Brain Heart Infusion broth (BHI; Oxoid) at 30 °C ± 1 °C for 24 h. Lactic bacteria were grown in Man, Rogosa and Shape broth (MRS; Oxoid) at 37 °C ± 1 °C for 24 h. Yeast and mold were cultivated in Malt Extract Broth (MEB, Oxoid) for at 25 °C ± 1 °C respectively for 72 h and 5 days.

2.3. Direct streaking

Direct streaking was performed by streaking 10 µL of the food item in duplicate on one agar plate. Only if there was colonies growth on both streaks was a result considered positive. For low acid food items (pH > 4.5), the direct streaking was performed on Plate Count Agar plates (PCA; Oxoid) and incubated at 30 °C ± 1 °C for 72 h ± 3 h. For high acid food items (pH < 4.5), the direct streaking was performed on Orange Serum Agar (OSA; Oxoid) plates incubated at 30 °C ± 1 °C for 3, 5 and 7 days. For high acid food items, the pH of the OSA agar (pH 5.5) was adjusted with hydrochloric acid (10%) to the pH of the food item. Direct streaking was performed in parallel with all the methods. Additionally, post-incubation streaking from vials of the Greenlight, the Bact/Alert and the Soleris systems was performed in this study to confirm the results.

2.4. Commercial alternative methods

A total of six commercial alternative methods were tested. Three methods (Greenlight, Bact/Alert and Soleris) were based on the detection of metabolites during microorganism growth, two methods (MLS and Cellscan) were based on microbial ATP detection, and one method (Bactiflow) based on flow cytometry technology. The Greenlight system (Mocon, Minneapolis, USA) measures the consumption of O_2 during microorganism growth. The Bact/Alert (bioMérieux, Marcy l'Etoile, France) and the Soleris (Neogen, Lansing, USA) systems measure microbial growth in food matrices based on CO_2 production. The flow cytometry method tested was the Bactiflow (bioMérieux, Marcy l'Etoile, France) where the microbial growth was measured in food matrices based on cell labelling and subsequent cell counting. The two ATP-based methods were the MLS system (3M, Minnesota, US) and the Cellscan system (Celsis, Chicago, US). The principle of these methods is based on the measurement of light emission produced due to the presence of ATP, which is involved in an enzyme substrate reaction between luciferase and luciferin. Free ATP present in the food item was first eliminated by addition of an ATP-ase. For all methods, the threshold was defined based on supplier instructions.

2.5. Microorganism enumeration

Aerobic spoilage microorganisms were enumerated on PCA (Oxoid) and incubated at 30 °C for 72 h ± 3 h. Lactic acid bacteria were enumerated on MRS (Oxoid) agar and incubated at 30 °C for 72 h ± 3 h. For yeast and mold, the enumeration was performed on Dichloran Glycerol agar base 18 (DG 18; Oxoid) and incubated at 25 °C for 5 days. Dilutions −4 to −8 were plated to perform the enumeration.

2.6. Inclusivity

A panel of 28 microorganisms (bacteria, yeasts and molds) capable to spoil the tested food items were selected based on literature and internal communications (Table 3). Cells were first grown under optimum conditions (2.2) and were subsequently diluted to 10^3–10^4 CFU/mL for Greenlight, Bact/Alert and Soleris systems, 10^4–10^5 CFU/mL for Bactiflow system and 10^6–10^7 CFU/mL for ATP-based methods.

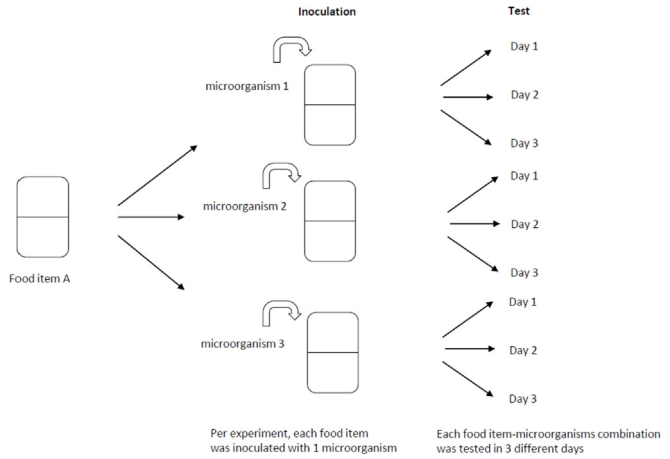

Fig. 1. Food item inoculation scheme. Each food item was tested with three different microorganisms. The inoculation was performed with one microorganism at a time. Each food item-microorganism combination was analyzed on three different days giving three sets of data.

2.7. Limit of detection (LOD_{95})

2.7.1. Food items inoculation

Each food item was challenged by three different microorganisms; at least one sporeforming microorganisms (*Bacillus*) to represent a contamination due to heat treatment failure and two non-sporeforming microorganisms to illustrate post-process contamination (Fig. 1). No spores were used as the validation protocol did not include the product incubation step and assumed that microorganisms were recovered from stress and were in the vegetative state. Details of the inoculation are shown in Fig. 2.

2.7.2. Analytical procedures

For the Greenlight system, either TSB or TSB for which the pH was adjusted with hydrochloric acid (10%) were used. For the other methods, proprietary media were used and tests were performed according to suppliers' instruction. For the Greenlight, the Bact/Alert and the Soleris systems, the test duration was set for 72 h at 30 °C ± 1 °C.

2.7.3. LOD_{95} determination

LOD_{95} were calculated for each method, food item and microorganism. The design was chosen to allow the calculation of this limit using regression between the known $\log_{10}$ concentration of the microorganism and a method outcome variable (or $\log_{10}$ outcome variable). For the Greenlight, the Bact/Alert and the Soleris systems the outcome variable was chosen as the time to detection (TTD) and the assessed cut-off values on this outcome variable was chosen as 72 h. In addition, 48 h and 24 h were also assessed in order to evaluate the gain of performance linked to the time to result. For these methods, linear regression between TTD and $\log_{10}$ concentration was used (Fig. 3). For MLS, Cellscan and Bactiflow, a better fit was obtained by using quadratic regression between the outcome variable and $\log_{10}$ concentration (Fig. 4). The assessed cut-off values were chosen as the method threshold. For each regression, the LOD_{95} was defined as the concentration corresponding to the intersection between the lower 95% confidence limit of prediction and the cut-off value. For direct streaking there is no continuous outcome, thus it was not possible to apply the same approach. The upper tolerance limit at 95% of the lower concentration for which the method gave a positive result was defined as LOD_{95}.

3. Results and discussion

3.1. Inclusivity

In contrast to ISO 16140–2 where inclusivity is not mandatory for large group of microorganisms such as total viable count or yeast and mold, for commercial sterility methods assessing inclusivity is imperative, but strains must be relevant to the selected food items. For this purpose, 28 microorganisms (Table 3) were selected based on literature and internal communication. Because of the principle of the tested methods, only aerobic microorganism were tested. Six microorganisms were not detected by one or several methods. *Geobacillus stearothermophilus*, mainly a concern for the dairy industry (Kakagianni, Gougouli, & Koutsoumanis, 2016) did not grow consistently in our trials, thus the results were not taken into consideration. *Bacillus coagulans*, an acidic-thermotolerant microorganism implicated in several tomato based products spoilage (Rice & Pederson, 1954), was detected by all the methods except the Bact/Alert system. *Micrococcus luteus* was not detected by the Soleris system and post-incubation streaking result was negative, indicating no growth occurred during the test. This microorganism should grow in a standard enrichment broth as it is an obligate aerobe. Its detection by Bact/Alert and Greenlight systems suggests that the growth failure could be linked to the lack of oxygenation of the broth. Bact/Alert system has a vial-shaking system incorporated in the incubation unit allowing a high oxygenation of the media that could be beneficial for the growth of *M. luteus*. The Greenlight system is based on the consumption of dissolved O_2 in the medium

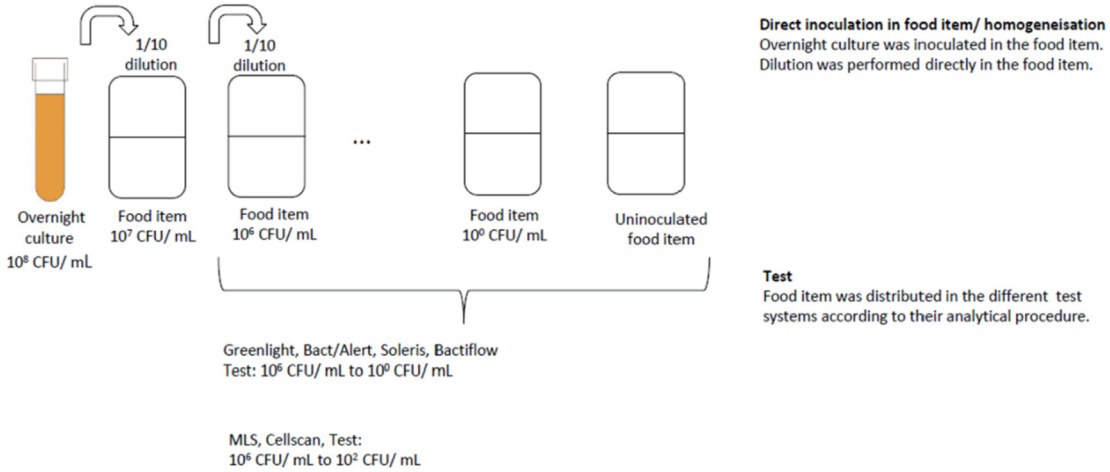

Fig. 2. Food item inoculation protocol. Overnight culture was directly inoculated in the food item at different concentrations. The inoculated food items were then distributed in the test systems.

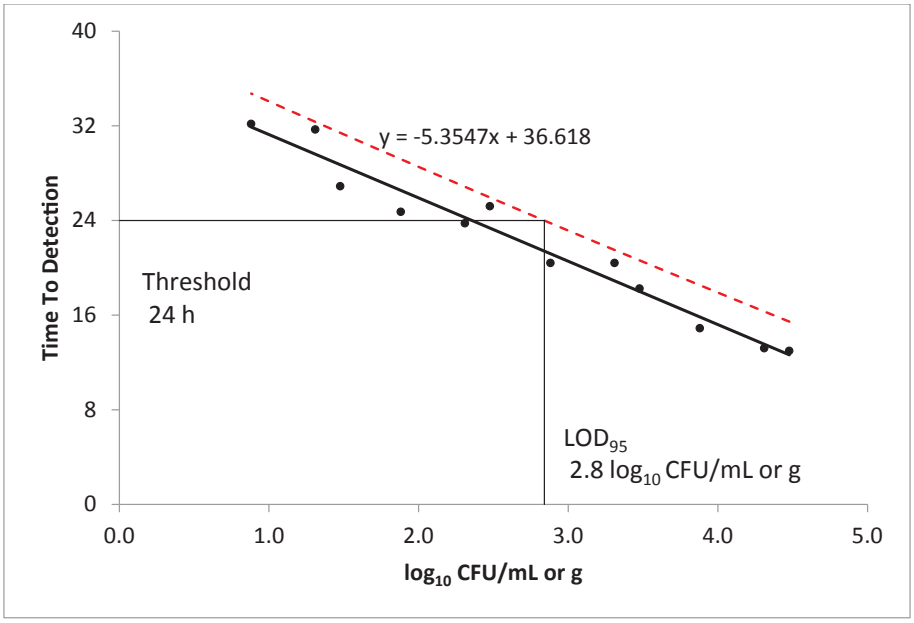

Fig. 3. Determination of LOD_{95} of the Greenlight, Bact/Alert and Soleris systems
egi104FK67D42S Measured data point egi10MR533V5ST Linear regression on data points egi10WW9BJ821K 95% upper confidence interval of prediction.

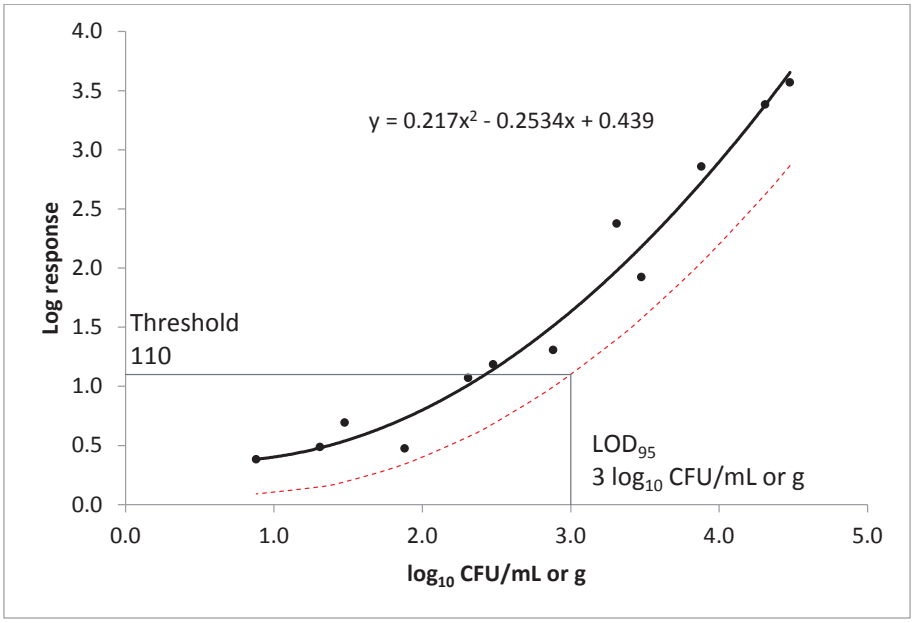

Fig. 4. Determination of LOD_{95} of the Bactiflow and ATP based methods
egi104FK67D42S Measured data point egi10MR533V5ST Linear regression on data points egi10WW9BJ821K 95% upper confidence interval of prediction.

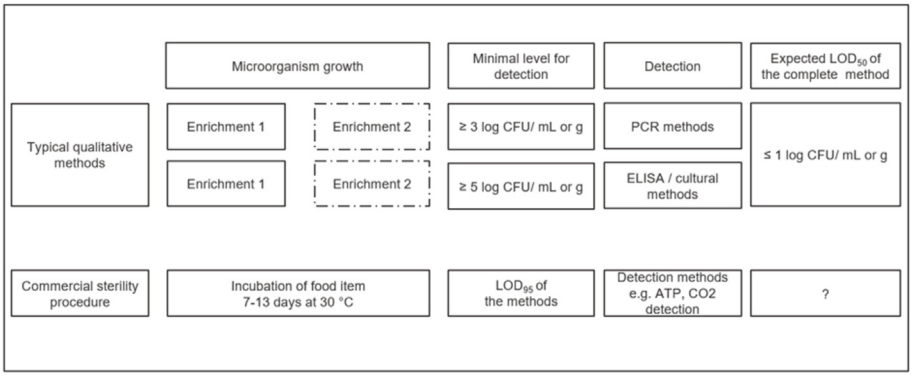

Fig. 5. Limit of detection of traditional detection methods against commercial sterility methods. Typical qualitative methods include at least one enrichment step to grow the microorganism and target a low limit of detection. A second enrichment tends to be optional (dashed outline) due to technologies with increasing sensitivity.

Table 4
LOD_{95} results.

Food item	Strain	LOD_{95} (log_{10} CFU/mL)												
		Greenlight			Bact/Alert			Soleris			MLS	Cellscan	Bactiflow	Direct streaking
		24 h	48 h	72 h	24 h	48 h	72 h	24 h	48 h	72 h	–	–	–	3 to 7 days[d]
Non-hydrolysed infant formula pH 6.77	*B. licheniformis*	< 1	< 1	< 1	2.0	< 1	< 1	< 1	< 1	< 1	5.3	5.3	> 6.3	4.24
	M. luteus	1.2	< 1	< 1	2.3	< 1	< 1	c	c	c	3.7	3.0	> 4.6	4.69
	L. plantarum	5.7	< 1	< 1	3.9	< 1	< 1	< 1	< 1	< 1	5.4	5.1	3.3	4.05
Hydrolysed infant formula pH 6.71	*B. licheniformis*	< 1	< 1	< 1	1.7	< 1	< 1	< 1	< 1	< 1	5.5	4.7	3.5	4.07
	P. fluorescens	< 1	< 1	< 1	1.2	< 1	< 1	< 1	< 1	< 1	6.2	5.2	4.5	4.31
	B. subtilis	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	6.1	4.8	3.9	4.71
High protein/vanilla flavor drink pH 6.89	*B. licheniformis*	< 1	< 1	< 1	1.6	< 1	< 1	< 1	< 1	< 1	6.0	4.4	4.4	2.7
	P. fluorescens	< 1	< 1	< 1	1.8	< 1	< 1	< 1	< 1	< 1	> 7.4	5.8	7.7	2.7
	S. warneri	< 1	< 1	< 1	3.4	< 1	< 1	2.6	< 1	< 1	6.1	4.6	3.3	3.4
Evaporated milk pH 6.40	*B. licheniformis*	< 1	< 1	< 1	2.8	< 1	< 1	1.7	< 1	< 1	5.9	6.1	3.3	2.3
	E. cloacae	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	7.1	5.9	4.6	3.7
	S. warneri	< 1	< 1	< 1	2.5	< 1	< 1	< 1	< 1	< 1	6.4	6.0	4.4	4.6
Chocolate dairy pH 6.75	*B. licheniformis*	< 1	< 1	< 1	2.4	< 1	< 1	< 1	< 1	< 1	6.3	6.2	4.6	4.0
	P. fluorescens	< 1	< 1	< 1	1.6	< 1	< 1	< 1	< 1	< 1	> 7.4	6.1	5.4	3.2
	E. cloacae	< 1	< 1	< 1	< 1	< 1	< 1	a	a	a	7.0	5.9	3.6	3.5
Banana-apple-strawberry puree pH 3.9	*Z. bailii*	5.3	3.6	2.2	2.8	< 1	< 1	2.1	< 1	< 1	a	a	> 5.4	3.0
	S. cerevisiae	4.2	2.4	< 1	2.7	< 1	< 1	< 1	< 1	< 1	4.4	4.1	3.7	2.5
	A. carbonarius	a	a	a	a	a	a	a	a	a	a	a	> 3.1	3.0
High energetic drink pH 4.13	*L. brevis*	4.4	< 1	< 1	2.8	< 1	< 1	3.3	< 1	< 1	8.0	5.7	3.6	> 12
	L. sphaericus	c	c	c	c	c	c	c	c	c	6.3	4.7	> 5.2	> 7
	S. cerevisiae	3.1	< 1	< 1	1.9	< 1	< 1	a	a	a	> 6.7	4.7	3.5	3.9
Carrot puree pH 5.1	*E. cloacae*	< 1	< 1	< 1	1.7	< 1	< 1	< 1	< 1	< 1	6.6	5.9	3.8	3.6
	Z. bailii	3.8	2.2	< 1	a	a	a	a	a	a	4.6	> 5.4	4	3.0
	B. pumilus	< 1	< 1	< 1	1.8	< 1	< 1	a	a	a	6.9	6.7	4.5	4.0
Béchamel sauce pH 6.53	*B. subtilis*	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1	a	4.3	4.1	2.8
	S. warneri	1.0	< 1	< 1	3.4	< 1	< 1	1.4	< 1	< 1	6.4	4.4	4.3	3.8
	L. plantarum	3.3	< 1	< 1	3.0	< 1	< 1	3.8	< 1	< 1	6.6	4.9	4.0	4.7

a: not consistent results to determine a LOD_{95} value.
b: not enough data points.
c: no growth.
d: low acid food items (pH > 4.5) were plated on PCA and incubated up to 3 days. High acid food item (pH < 4.5) are were plated on OSA and incubated up to 7 days.

and the detection of *M. luteus* could have been linked by the high sensitivity of the O_2 sensor present in the analytical tube. Previous studies performed within Neogen laboratories indicated indeed that some *M. luteus* strains were not detected by the Soleris system (internal communications). Additional work with this *M. luteus* strain and other obligate aerobes would provide insight into the detection of obligate aerobes with the Soleris system. *Eurotium amstelodami* was not detected by the Soleris system nor by the two ATP-based methods. *E. amstelodami* is a xerophilic mold that grows better at low a_w media (a_w 0.95) (Hocking & Pitt, 1980). Possible reason is that MEB broth used to activate the strain did not support its growth to the level needed for the detection with ATP-based methods and the broth (OSB) used in Soleris protocol was not suitable. *Leuconostoc mesenteroides* is a spoilage microorganism associated with several food types including fruit and vegetable (De Paula, Jeronymo-Cenevita, Todorov, & Penna, 2015) which could potentially spoil food items such as purees. It was not detected by the Greenlight and the MLS systems. In the case of Greenlight, the negative result was likely due to the use of TSB at pH 4 as specified in the protocol (2.7.2). The optimum pH growth for *Leuconostocs* lies between pH 4 and 6, but for some specific strains, growth may be affected by a pH lower than 4.5 (Mataragas, Metaxopoulos, Galiotou, & Drosinos, 2003). In addition to *L. mesenteroides,* the MLS system did not detect *Aspergillus carbonarius*. The two microorganisms were grown in identical conditions for the ATP-based methods MLS and Cellscan. However, only the latter correctly detected both organisms. The observed negative result of MLS system was not understood, further tests with lactic bacteria and molds would be required.

3.2. Assessment of method sensitivity

The design proposed here, allows to assess the performance of each method as an absolute value and not in comparison with the performance of a reference method as described in ISO 16140–2. It aims to establish the sensitivity of each method after product incubation step (Fig. 1). The most relevant performance criterion was LOD_{95}, a lowest microorganism concentration, when present, allows to detect the microorganism in 95% of the cases. In addition, the selected methods provide a quantitative value (TTD, RLU, cells count) associated to the microorganism concentration in a food item. The use of quantitative values as an outcome allows to calculate the LOD_{95} with a limited number of measures in high level of confidence. To determine the LOD_{95}, two ranges were used. The first, a low range (1 CFU - 6 log_{10} CFU/mL), was used for Greenlight, Bact/Alert and Soleris systems. A low LOD_{95} was expected, since these methods include in its protocol an incubation step that allows measuring cellular metabolisms during microbial growth. In contrast, a higher inoculum range (2–6 log_{10} CFU/mL) was used for the Bactiflow and the two ATP-based methods since these methods do not include an incubation step. In most cases, the regression provided a good correlation (R^2 around 0.9) between the concentration of the microorganism (log_{10} CFU/mL) and the alternative methods response (TTD, RLU, cells count).

3.2.1. Detection limit of the direct streaking method

The protocol described in previous section allows to assess the performance of the alternative methods, but not to determine an accurate LOD_{95} of the direct streaking method. Indeed, contrary to the alternative methods, the outcome of the direct streaking method is a qualitative value (presence/absence). For most experiments, the lowest

concentration at which the direct streaking method gave a positive result was higher or equal to the actual limit of detection. Furthermore, since the inoculation level was calculated using a $\log_{10}$ cfu/mL scale, the actual limit of detection could fall within a range of one $\log_{10}$ cfu/mL of the reported value. There are no references describing the detection limit of the direct streaking method. However, a theoretical limit can be determined since 3 $\log_{10}$ CFU/mL of microorganisms would have to be present to obtain at least one colony on plate using a 10 μL loop. The LOD_{95} values obtained in this study were between 2.2 $\log_{10}$ CFU to 4.7 $\log_{10}$ CFU/mL indicating the actual values could be higher than the theoretical value and it showed that the limit of detection would depend on the strain-food item combination (Table 4). Moreover, on two occasions, no colonies were observed. In one of these cases (*Lysinibacillus sphaericus* in high energetic drink at pH 4.13) the absence of colonies on the plate could be explained by the inhibition of the microorganism at pH 4.1. This result aligned with *L. sphaericus* growth characteristics described in literature (Ahmed, Yokota, Yamazoe, & Fujiwara, 2007). In the second case (*Lactobacillus brevis* in high energetic drink), the absence of colonies on plate was not linked to the absence of growth in the food item as *L. brevis* is normally not inhibited at pH 4. In addition, all the alternative methods were able to detect the microorganism. Further work is needed to understand this false negative result.

3.2.2. LOD_{95} of the greenlight, the bact/alert and the soleris systems

For the three methods based on the measurement of metabolites, the LOD_{95} values were less than 1 $\log_{10}$ CFU/mL for dairy food items after 72 h of test (Table 4). These results were expected as dairy based food items offer a nutritious and stress-free environment for bacteria. Moreover, the related spoilage microorganisms are not considered fastidious. For more challenging food items (food item with pH $<$ 4.5), the LOD_{95} values were most of the cases below 1 $\log_{10}$ CFU/mL. However, there were some exceptions. On one occasion (*Zygosaccharomyces bailii* in banana-apple-strawberry puree) values of 2.2 $\log_{10}$ CFU/mL was obtained by the Greenlight system, which was still below the theoretical value of the direct streaking method. For four food item and strain combinations (*E. cloacae* in chocolate dairy, *A. carbonarius* in banana-apple-strawberry puree, *Zygosaccharomyces bailii* in carrot puree and *Bacillus pumilus* in carrot puree), one or several methods did not provide consistent results to draw a conclusion. Finally, *L. sphaericus* in high energetic drink was not detected by any method confirming the inhibition of this strain at pH 4.13 (3.2.1). *M. luteus* was not detected by the Soleris system, confirming the inclusivity result. These limitations indicates that an adaptation in the analytical protocol (media, dilution factor) or optimization of software parameters is required. For the commercial sterility methods that have an incubation step, the LOD_{95} is generally very low ($<$ 1 $\log_{10}$ CFU/mL), thus emphasizing their high sensitivity and questioning whether the incubation time of the food items prior testing can indeed be reduced. On the other hand, because of this high sensitivity, media used in the analytical protocol should reflect the conditions of the food item in order to avoid unexpected growth that give false positive results. In addition to the validation of the standard analytical protocol (test duration for 72 h), the LOD_{95} design allowed to evaluate the performance of the methods at shorter test durations (24 and 48 h). For dairy based food items, the results with a test duration of 48 h were identical to 72 h, indicating a reduction of the test duration to 48 h was possible for those food items. In contrast, a test duration of 48 h for challenging food items and 24 h for all tested food items gave higher LOD_{95} (values up to 5 $\log_{10}$ CFU/mL), indicating in these cases, a considerable decrease of the methods sensitivity.

3.2.3. LOD_{95} of the bactiflow and the ATP-based methods

Contrary to the previously described methods, the Bactiflow system and the ATP-based methods do not include an incubation step, but a method response is given directly either by the cell count or the ATP value thus require a higher level of contamination. Indeed, the LOD_{95} values of these alternative methods were higher, ranging from 3 $\log_{10}$ CFU to more than 7 $\log_{10}$ CFU/mL (Table 4). In addition to the high LOD_{95} values, a number of strain and food item combinations were not detected by one or several methods. On three occasions, the MLS system did not detect the contamination (*P. fluorescens* in high protein drink, *P. fluorescens* in chocolate dairy and *S. cerevisiae* in high energetic drink) and in three other occasions, the results were inconsistent to draw a conclusion (*B. subtilis* in béchamel sauce, *Z. bailii* and *A. carbonarius* in banana-apple-strawberry puree). The Cellscan system did not detect the contamination in one case (*Z. bailii* in carrot puree) and generated inconsistent results in two other cases (*Z. bailii* and *A. carbonarius* in banana-apple-strawberry puree). The Bactiflow system was not able to detect the contamination on five occasions (*B. licheniformis* and *M. luteus* in non-hydrolysed infant formula, *Z. bailii* and *A. carbonarius* in banana-apple-strawberry puree and *L. sphaericus* in high energetic drink). While the failure to detect yeast and mold in high acid food items may be related to a non-optimal analytical protocol, the negative results obtained with bacterial cultures could not be explained. It should be mentioned that only the cut-off value that was defined in the method instructions was used and a different cut-off value is likely to have an impact on the method sensitivity. With the protocol described in this paper, the LOD_{95} showed that the Bactiflow and the ATP-based methods were able to detect spoilage microorganisms only at high concentrations (3 $\log_{10}$ CFU to 7 $\log_{10}$ CFU/mL).

3.2.4. Acceptance criterion of LOD_{95}

In a method validation protocol, defining the acceptance criteria is as important as the validation experimental design itself. In a typical microbiological detection method, a food item is first incubated in an enrichment broth to allow the target microorganism to grow and subsequently tested by a specific technology e.g. PCR, ELISA. Optimized sample preparation have been also developed (e.g. ISO 6887 series) to guarantee the good growth of the target microorganism present at low level in the food item. The objective of the enrichment step is to bring the microorganism concentration to the limit of detection of the device (e.g. $\geq$ 3 $\log_{10}$ CFU/mL for PCR methods, $\geq$ 5 $\log_{10}$ CFU/mL for ELISA methods). The limit of detection of the complete method is therefore as low as possible and an acceptance criterion is defined based on this value (e.g. $\leq$ 1 $\log_{10}$ CFU/mL). In commercial sterility, there is no enrichment broth, but the food item is incubated between seven and thirteen days prior to testing. Only microorganisms that can grow in the food item with its intrinsic hurdles such as low pH and fat content are targeted. Bearing this in mind, the expected microorganism concentration is unknown and consequently the acceptance criterion is not easy to define (Fig. 5).

Based on the theoretical value of the direct streaking method, the acceptance criteria would be 3 $\log_{10}$ CFU/mL or g. However as it showed in 3.2.1, positive direct streaking results are much depended on the strain-food item combination. For that reason, we have chosen to base the absolute acceptance value on the expected level at the end of the product incubation (7–13 days at 30 °C). The acceptance criterion was thus linked to the maximum growth (*N* max) of the spoilage microorganism in the food item. Literature shows that for low acid food items, the *N* max after 10 days was 8 $\log_{10}$ CFU/mL and for high acid food items, the *N* max was 6 $\log_{10}$ CFU/mL using *Bacillus cereus* as a model (Soares, Kabuki, & Kuaye, 2012; Wong, Chen, & Chen, 1988). Based on these values, the suggested acceptance criterion was $LOD_{95} \leq$ 6 $\log_{10}$ CFU/mL.

3.3. Inter-laboratory study

In this paper, we proposed a validation design for a single laboratory. However, most validation guidelines would include an inter-laboratory study as part of the validation design. The inter-laboratory study is to determine the sensitivity of the method when executed by

different laboratories using the same food item - strain combination and gives an indication whether the method can be successfully implemented in other laboratories. The design of this study would be based on the qualitative determination (detected/not detected) of the LOD_{95} value determined during the single laboratory study. To achieve this, three levels of inoculation in a specific food item would be prepared and send to the laboratories. The three levels would be one negative level (a), one corresponding to the LOD_{95} determined during the single laboratory study (b) and one level slightly above the LOD_{95} (c). The acceptance criterion for the inter-laboratory study would be a positive detection of the level above the LOD_{95} (c) for all participating laboratories. Indeed, due to LOD_{95} uncertainty, there is still the possibility that a sample inoculated at the LOD_{95} level (b) may not be detected by few laboratories then the higher LOD_{95} level (c) should be chosen to overcome this possibility. ISO 16140-2 describes in more detail the study logistic and required number of laboratories.

3.4. ISO 16140 series

The European Regulation EC 2073/2005 and the related amendment 2019/229 recognize the use of reference methods or the use of alternative methods if their performance is equivalent to the reference method. ISO 16140 series has been developed to propose a common reference protocol to validate qualitative and quantitative methods, covering a large range of microbiological methods. However, there are still methods that require specific validation protocols e.g. non-culturable microorganisms and parasites. The application of innovative technologies may also require additional validation protocols or the revision of ISO 16140–2. From that perspective, the protocol described here could be proposed as an additional part in the ISO 16140 series.

4. Conclusion

Before implementing an alternative method, a food business operator should demonstrate that the method is fit-for-purpose; method validation is therefore a key process. To date, commercial sterility methods are not validated since existing protocols cannot be applied as such. In this paper, we describe a protocol that includes an inclusivity and LOD_{95} performance criteria to assess method selectivity and the sensitivity. The inclusivity criterion allows to determine whether a method can detect the relevant microorganisms. The LOD_{95} criterion permits to define the lowest contamination level at which the microorganisms are detected in 95% of cases. The inclusivity results showed that the selection of the relevant microorganisms provided an important insight on the performance of the method, either from the perspective of the analytical protocol or the technology. The LOD_{95} results showed that methods based on measurement of metabolic activity during an incubation step were very sensitive (LOD_{95} of < 1 $\log_{10}$ CFU/mL). This low detection limit suggests that; (1) the test duration could be decreased and/or (2) the product incubation time prior to analysis could be reduced. These methods present a promising alternative to the traditional direct streaking. Nevertheless, an optimization of the analytical protocol (sample preparation, software parameters) is still required for some complex food items and media used must reflect the food item conditions to avoid false positives. The flow cytometer and ATP-based methods provided a faster results (hours instead of days), but their disadvantage is that they are less sensitive ($LOD_{95} > 3$ $\log_{10}$ CFU/mL). Also here an optimization of the analytical protocol and cut-off value could increase the sensitivity. The current high LOD_{95} values suggest that for those methods, the product incubation time of 7–13 days should be maintained to avoid false negative results. The acceptance criterion is based on the estimated growth after product incubation and the proposed value is $LOD_{95} \leq 6$ $\log_{10}$ CFU/mL. Furthermore, our protocol including the inter-laboratory study could be a proposal for the ISO 16140 series. Finally, with a protocol focusing only on the detection step that provides the sensitivity of the available methods, optimization of the product incubation time (e.g. reduce the incubation time) could be performed independently. The results of this validation protocol will give food business operators and method suppliers a valuable insight into the performance of the methods, and provide assurance that the obtained results are valid and meet their specifications and requirements.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors would like to thank Prof Mieke Uyttendaele for valuable comments on the manuscript and Dr Mariem Ellouze for inputs on microbial growth kinetics.

References

Ahmed, I., Yokota, A., Yamazoe, A., & Fujiwara, T. (2007). Proposal of *Lysinibacillus boronitolerans* gen. nov. sp. nov., and transfer of *Bacillus fusiformis* to *Lysinibacillus fusiformis* comb. nov. and *Bacillus sphaericus* to *Lysinibacillus sphaericus* comb. nov. *International Journal of Systematic and Evolutionary Microbiology, 57*, 1117–1125.

Anonymous (1993). Code of Hygienic practices for aseptically processed and packaged low-acid foods. *CAC/RCP 40*.

Anonymous (1997). Microbiologie des aliments - Contrôle de la stabilité des produits appertisés et assimilés - Méthode de référence. *Association Française de Normalisation (AFNOR)* (France).

Anonymous (2016). *ISO 16140-2: 2016. Microbiology of the food chain- method validation- Part 2: Protocol for the validation of alternative (proprietary) methods against a reference method.* Geneva, Switzerland: Int. Org. for Standardization.

Bolzoni, G., Marcolini, A., Delle Donne, G., Appicciafuoco, B., & Ferrini, A. M. (2015). New national conversion line for bactoscan FC in Italy: A step forward. *Italian Journal of Food Science, 27*, 191–197.

Datta, N. (2018). UHT treatment of milk and dairy products. *Advances in Dairy Research, 6*, 1.

De Paula, A. T., Jeronymo-Cenevita, A. B., Todorov, S. D., & Penna, A. L. B. (2015). The two faces of *Leuconostoc mesenteroides* in food systems. *Food Reviews International, 31*, 147–171.

Eneroth, A., Ahrneh, S., & Molin, G. (2000). Contamination routes of Gram-negative spoilage bacteria in the production of pasteurized milk, evaluated by randomly amplified polymorphic DNA (RAPD). *International Dairy Journal, 10*, 325–331.

Feldsine, P., Abeyta, C., & Andrews, W. H. (2002). AOAC INTERNATIONAL methods Committee guidelines for validation of qualitative and quantitative food microbiological official methods of analysis. *Journal of AOAC International, 85*, 1187–1200.

Fernandes, R., Carey, C., Hynes, J., & Papkovsky, D. (2013). GreenLight™ model 960. *Journal of AOAC International, 96*, 369–385.

Hocking, A.,D., & Pitt, J. I. (1980). Dichloran-glycerol medium for enumeration of xerophilic fungi from low-moisture foods. *Applied and Environmental Microbiology, 39*, 488–492.

Kakagianni, M., Gougouli, M., & Koutsoumanis, K. P. (2016). Development and application of *Geobacillus stearothermophilus* growth model for predicting spoilage of evaporated milk. *Food Microbiology, 57*, 28–35.

Krumm, G. W., Lattuada, C. P., Johnston, R. W., Eye, J. G., & Green, J. (1998). *Examination of heat processed, hermetically sealed (canned) meat and poultry.* USDA/FSIS Microbiology Laboratory Guidebook (Chapter 10).

Landry, W. L., Schwab, A. H., & Lancette, G. A. (2001). Examination of canned foods. *Bacteriological Analytical Manual* (Chapter 21(A)).

Mataragas, M., Metaxopoulos, J., Galiotou, M., & Drosinos, E. H. (2003). Influence of pH and temperature on growth and bacteriocin production by *Leuconostoc mesenteroides* L124 and *Lactobacillus curvatus* L442. *Meat Science, 64*, 265–271.

Mostert, F. J., & Jooste, P. J. (2005). *Dairy microbiology handbook: The microbiology of milk and milk products* (3rd ed.). .

Mozola, M., Gray, L. R., Feldpausch, J., Alles, S., McDougal, S., Montei, C., et al. (2013). Validation of the Soleris® NF-tvc method for determination of total viable count in a variety of foods. *Journal of AOAC International, 96*, 399–403.

Pujol, L., Albert, I., Johnson, N. B., & Membré, J. M. (2013). Potential application of quantitative microbiological risk assessment techniques to an aseptic-UHT process in the food industry. *International Journal of Food Microbiology, 162*, 283–296.

Rice, A. C., & Pederson, C. S. (1954). Factors influencing growth of *Bacillus coagulans* in canned tomato juice. I. size of inoculum and oxygen concentration. *Journal of Food Science, 19*, 115–123.

Ropkins, K., & Beck, A. J. (2000). Evaluation of worldwide approaches to the use of HACCP to control food safety. *Trends in Food Science & Technology, 11*, 10–21.

Sanchez-Madrid, M. F. (2003). *Preservation of food. Encyclopedia of food science and nutrition* (2nd ed.). .

Soares, C. M., Kabuki, D. Y., & Kuaye, A. Y. (2012). Growth of enterotoxin producing *Bacillus cereus* in meat substrate at 10 °C and 30 °C. *Brazilian Journal of Microbiology,* 1401–1405.

Sperber, W. H. (2005). HACCP does not work from farm to table. *Food Control, 16*, 511–514.

Wong, H. C., Chen, Y. L., & Chen, C. L. F. (1988). Growth, germination and toxigenic activity of Bacillus cereus in milk products. *Journal of Food Protection, 51*, 707 701.